

# SESTAT: A TOOL FOR STUDYING SCIENTISTS AND ENGINEERS IN THE UNITED STATES

Within the U.S. population exists a group of individuals generically referred to as scientists and engineers. This group includes both people educated in science and engineering (S&E) fields and individuals who, although not educated in these fields, are working in S&E occupations. They include among their number technicians and technologists, researchers, educators, and managers of the S&E enterprise. Although these workers constitute only a small fraction of the total U.S. labor force, their impact on our society exceeds their number. Scientists and engineers contribute to technological innovation and economic growth, to scientific and engineering research, and to a greater understanding of S&E.

Determining the population of scientists and engineers in the U.S. can be dramatically affected by the criteria used to define this group. Educational degree levels and fields, occupational categories, or a combination of these factors may all be taken into account. To better define and understand this population, the National Science Foundation (NSF) developed the Scientists and Engineers Statistical Data System, or SESTAT ('*SEE-stat*').

SESTAT was designed in response to recommendations of the National Research Council's Committee on National Statistics (CNSTAT). The committee, with support from NSF, was asked in the late 1980s to conduct a comprehensive review of the NSF's surveys and data on scientists and engineers and to propose methods and procedures for increasing the quality and usefulness of the data in the 1990s. As NSF was expecting to draw a new sample of experienced scientists and engineers from the 1990 decennial census, the opportunity was provided to examine the design and operation of NSF's S&E data system. The committee's analyses and recommendations for the data system were presented in their 1989 report Surveying the Nation's Scientists and Engineers: A Data System for the 1990s. As one of its principal proposals, CNSTAT stated:

*"We strongly urge that the NSF personnel data system for the 1990s strive to provide information that will permit users to apply their own definitions of the science and engineering population to suit their particular research and analysis purposes within a framework that*

*facilitates cross-comparisons with other widely used data sources. Specifically, we believe that the system should support analysis of the science and engineering community from each of two major perspectives... from the perspective of occupational employment or jobs and from the perspective of academic training or careers."*

The NSF has strongly endorsed this recommendation and it has served as the guiding principle in the design and development of SESTAT.

## WHAT IS SESTAT?

SESTAT is a comprehensive and integrated system of information about the employment, educational, and demographic characteristics of scientists and engineers in the United States. It comprises data collected through three national sample surveys supported by NSF: the National Survey of College Graduates (NSCG), the National Survey of Recent College Graduates (NSRCG), and the Survey of Doctorate Recipients (SDR). These surveys are conducted biennially; each is administered to a different sample population of bachelor's and above college degree holders. Although there are individuals in the U.S. who have earned science and engineering related degrees at the associate's degree level or who do not have bachelor's degrees but are working in science and engineering occupations, these individuals are not included in the sampling frames for the three SESTAT surveys.

## THE NATIONAL SURVEY OF COLLEGE GRADUATES

The NSCG was first administered in 1993 and biennially thereafter to a nationally representative sample of all college degree holders who were identified through the 1990 decennial census. The target population for this survey includes individuals in the United States as of April 1990 with a bachelor's degree or higher in any field, not just the sciences or engineering. Besides capturing people with degrees earned at U.S. institutions, the NSCG also includes college degree holders who earned their degrees outside of the United States but who were residing here in 1990. In 1993, two selected groups from the NSCG were incorporated into the SESTAT database: those with science or engineering

Table 1. Who is represented in the SESTAT integrated database?			
Survey	1993 SESTAT	1995 SESTAT	1997 SESTAT
National Survey of College Graduates	Bachelor's level and above S&E degree holders as of April 1990, except those with U.S.-earned doctorates	Bachelor's level and above S&E degree holders as of academic year 1994, except those with U.S.-earned doctorates	Bachelor's level and above S&E degree holders as of academic year 1996, except those with U.S.-earned doctorates
	Bachelor's level and above non-S&E degree holders as of April 1990 who hold S&E jobs, except those with U.S.-earned doctorates	Bachelor's level and above non-S&E degree holders as of April 1990 who hold S&E jobs, except those with U.S.-earned doctorates	Bachelor's level and above non-S&E degree holders as of April 1990 who hold S&E jobs, except those with U.S.-earned doctorates
National Survey of Recent College Graduates	Individuals who earned bachelor's or master's S&E degrees in May to December of 1990, or academic years 1991 or 1992	Individuals who earned bachelor's or master's S&E degrees in academic years 1993 or 1994	Individuals who earned bachelor's or master's S&E degrees in academic years 1995 or 1996
Survey of Doctorate Recipients	Individuals who earned S&E doctorates in the U.S. through academic year 1992	Individuals who earned S&E doctorates in the U.S. through academic year 1994	Individuals who earned S&E doctorates in the U.S. through academic year 1996

**SOURCE:** National Science Foundation, Division of Science Resources Studies.

degrees, and those without such degrees but working in S&E occupations. These two populations are collectively referred to as the NSCG S&E panel.<sup>1</sup> In 1995 and subsequent rounds of the survey, these same two groups were followed.

## THE NATIONAL SURVEY OF RECENT COLLEGE GRADUATES

A second survey, the NSRCG – which has been administered biennially since the early 1970s to recent S&E bachelor's and master's degree recipients – is used to incorporate new U.S. S&E degree earners in the 1990s. In 1993, the NSRCG consisted of a sample of individuals who earned new S&E bachelor's and master's degrees in the 1990, 1991, and 1992 academic years. In 1995, these 1993 sample cases were moved into the NSCG sample frame; the 1995 NSRCG includes only recent U.S. S&E bachelor's and master's degree earners from the 1993 and 1994 academic years. The 1995 NSRCG sample was passed to the NSCG for the 1997 survey round; the 1997 NSRCG includes new U.S. S&E bachelor's and master's degree earners from the 1995 and 1996 academic years.

## THE SURVEY OF DOCTORATE RECIPIENTS

A third survey, the Survey of Doctorate Recipients, has been sponsored by the National Science Foundation and other

federal agencies since the early 1970s. In this survey, a sample of holders of S&E doctorates earned at U.S. institutions are followed throughout their careers from year of degree award until age 75. Every 2 years, a sample of new S&E doctoral degree earners is added to the SDR from another NSF-sponsored survey, the Survey of Earned Doctorates. In 1993, the SDR sample frame included all U.S.-earned S&E doctorates through academic year 1992; the 1995 sample frame includes doctorates earned through the 1994 academic year; the 1997 sample frame includes doctorates earned through the 1996 academic year.

## THE SESTAT INTEGRATED DATABASE

The SESTAT database was created by the integration of the S&E Panel of the NSCG, the NSRCG, and the SDR. Table 1 is a summary of the populations included in these three component surveys of the SESTAT integrated database. Because after 1993 the SESTAT surveys identify individuals for inclusion only at the point of earning a new S&E degree from a U.S. institution, two subpopulations of scientists and engineers in the United States are underrepresented in the SESTAT integrated database in 1995 and subsequent survey years: (1) new immigrants with S&E degrees earned outside the U.S. who entered the U.S. after 1990, and (2) people with no S&E degrees working in S&E occupations after 1990.

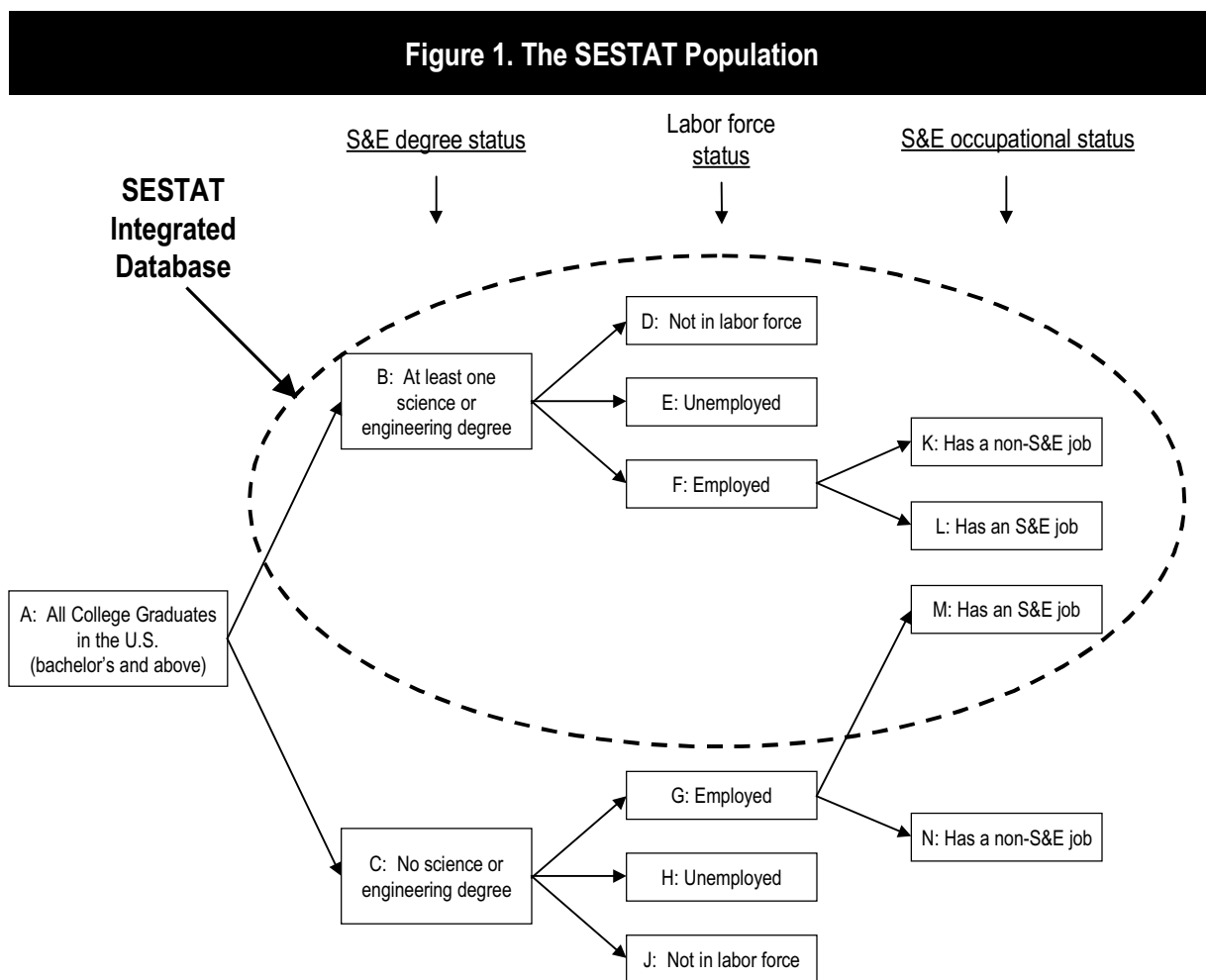
<sup>1</sup> For 1993, the full NSCG file (which includes all college degree holders, both S&E and non-S&E) is available for analysis. This file can be obtained on CD-ROM, or accessed through the SESTAT web site.

## HOW MANY PERSONS ARE INCLUDED IN SESTAT?

The SESTAT integrated database captures that part of the science and engineering population who either received a college degree (bachelor's or higher) in an S&E field or those who work in an S&E occupation with a bachelor's degree or higher in any field. This is a broad designation that includes many fields of study, as well as many occupations. The dashed oval on Figure 1 shows the groups of individuals included in the SESTAT integrated database. There are other groups in the U.S. population

that can be considered part of the population of scientists and engineers, for example persons in technical occupations without college degrees, who are not included in the SESTAT integrated database. The SESTAT system was developed to provide information on the bachelor's and above population of scientists and engineers, with emphasis placed on those trained at U.S. institutions.

The 1993 SESTAT integrated database represents 11,615,200 individuals. This included 11,021,500 persons with S&E degrees, and 593,600 persons without such degrees but working in S&E occupations (table 2).



**Note:** Because after 1993 the SESTAT surveys identify individuals for inclusion at the point of earning an S&E degree from a U.S. institution, two subpopulations of scientists and engineers in the United States are underrepresented in the SESTAT integrated database in 1995 and subsequent survey years: (1) new immigrants with S&E degrees earned outside the U.S. who entered the U.S. labor force after 1990, and (2) people with no S&E degrees working in S&E occupations after 1990 (Box M in the diagram). There are other individuals who could be considered part of the population of scientists and engineering, for example persons in technical occupations with two-year degrees or other types of training. These individuals are not included in the SESTAT integrated database, nor are they represented in this diagram.

**SOURCE:** National Science Foundation, Division of Science Resources Studies.

**Table 2: SESTAT integrated database population statistics**
**1993 SESTAT integrated database**

S&E degree status	Total	Labor Force				Not in labor force
		Employed			Unemployed	
		Total	In S&E	In non-S&E		
Total SESTAT population	11,615,200	9,793,500	3,303,400	6,490,100	322,200	1,499,500
Educated in S&E	11,021,500 <i>(Fig. 1, Box B)</i>	9,199,900 <i>(Fig. 1, Box F)</i>	2,709,800 <i>(Fig. 1, Box L)</i>	6,490,100 <i>(Fig. 1, Box K)</i>	322,200 <i>(Fig. 1, Box E)</i>	1,499,500 <i>(Fig. 1, Box D)</i>
Highest degree is in S&E	8,571,000	7,035,800	2,517,800	4,518,000	272,500	1,262,600
Highest degree is in non-S&E	2,450,600	2,164,100	192,000	1,972,100	49,600	236,900
No S&E degree	593,600	593,600	593,600 <i>(Fig. 1, Box M)</i>	0	0	0

**1995 SESTAT integrated database**

S&E degree status	Total	Labor Force				Not in labor force
		Employed			Unemployed	
		Total	In S&E	In non-S&E		
Total SESTAT population	12,036,200	10,114,500	3,185,600	6,928,900	249,300	1,672,400
Educated in S&E	11,456,600 <i>(Fig. 1, Box B)</i>	9,570,000 <i>(Fig. 1, Box F)</i>	2,851,400 <i>(Fig. 1, Box L)</i>	6,718,600 <i>(Fig. 1, Box K)</i>	237,700 <i>(Fig. 1, Box E)</i>	1,649,000 <i>(Fig. 1, Box D)</i>
Highest degree is in S&E	8,908,000	7,333,100	2,634,900	4,698,200	197,400	1,377,600
Highest degree is in non-S&E	2,548,600	2,236,900	216,600	2,020,400	40,300	271,400
No S&E degree*	579,600	544,400	334,100 <i>(Fig. 1, part of Box M)</i>	210,400 <i>(Fig. 1, part of Box N)</i>	11,600 <i>(Fig. 1, part of Box H)</i>	23,400 <i>(Fig. 1, part of Box J)</i>

\* The 245,400 persons without S&E degrees or jobs in the 1995 integrated database represent individuals who had S&E jobs in 1993, but by 1995 had moved to non-S&E jobs, become unemployed, or had moved out of the labor force. Although these cases are shown in the SESTAT integrated database, they no longer fit the population of individuals either educated in S&E or working in S&E. These individuals, along with the 334,100 persons in the 1995 integrated database without S&E degrees but still working in S&E jobs are part of a panel of individuals that are being followed throughout the 1990s. The latter group is described in figure 1 as "Box M", and is underrepresented in the SESTAT database after 1993 because the SESTAT surveys do not capture new persons entering S&E occupations who are not educated in S&E fields in this decade. The SESTAT surveys only cover persons who earned S&E degrees outside of the U.S. if they were residing in the country in 1990. Persons who earned S&E degrees outside of the U.S. and then entered after that time are not covered in the SESTAT integrated database.

**SOURCE:** National Science Foundation, Division of Science Resources Studies.

The 1995 SESTAT integrated database represents 12,036,200 individuals: 11,456,600 with S&E degrees; 334,100 persons without such degrees but working in S&E occupations, and 245,400 persons without S&E degrees or jobs.<sup>2</sup>

Approximately 84 percent of the individuals in the 1993 and 1995 SESTAT integrated databases (9,793,500 and 10,114,500, respectively) were employed. Among this group, about one-third were employed in science or engineering occupations (3,303,400 in 1993 and 3,185,600 in 1995). Individuals in this group are often referred to as those working in S&E. The remaining individuals were either unemployed, but seeking work (322,200 in 1993 and 249,300 in 1995) or were not in the labor force (1,499,500 in 1993 and 1,672,400 in 1995).

Many of the individuals in the 1993 and 1995 integrated databases hold either multiple S&E degrees or have degrees in both S&E and non-S&E fields. There were 11,021,500 individuals in 1993 who were educated in S&E; that number had risen to 11,456,600 by 1995. Approximately three-fourths of these individuals reported that their highest degree was in an S&E field (8,571,000 in 1993 and 8,908,000 in 1995).<sup>3</sup>

Many workers educated in S&E routinely find occupations in non-S&E fields. Roughly 70% of all S&E-educated individuals in the labor force were employed in non-S&E occupations (6,490,100 in 1993 and 6,718,600 in 1995). Conversely, but to a lesser extent, non-S&E educated individuals are employed in S&E occupations. In 1993, there were 593,600 non-S&E educated persons working in S&E occupations. In 1995, the SESTAT integrated database included only 334,100 persons in this category.<sup>4</sup>

---

<sup>2</sup> The 245,400 persons without S&E degrees or jobs in the 1995 integrated database represent individuals who had S&E jobs in 1993, but by 1995 had moved to non-S&E jobs, become unemployed, or had moved out of the labor force. Although these cases are shown in the SESTAT integrated database, they no longer fit the population of individuals either educated in S&E or working in S&E. These individuals, along with the 334,100 persons in the 1995 integrated database without S&E degrees but still working in S&E jobs are part of a panel of individuals that are being followed throughout the 1990s. The latter group is described in figure 1 as “Box M”, and is underrepresented in the SESTAT database after 1993 because the SESTAT surveys do not capture new persons entering S&E occupations who are not educated in S&E fields in this decade.

<sup>3</sup> For individuals with two or more degrees at the same highest degree level (bachelor’s, master’s, doctorate, professional), the field of the most recently earned degree is considered the highest degree field.

<sup>4</sup> See footnote 2.

## WHAT CAN WE FIND OUT ABOUT SCIENTISTS AND ENGINEERS USING SESTAT?

There is a wide range of information on scientists and engineers contained in SESTAT. Listed here are examples of the variables found in the SESTAT integrated database.

### Labor Force Information

For the employed:

- Primary job and salary
- If previously retired
- Type of employer: educational institution (by type); private for-profit; private not-for-profit; government (state/local or federal); self-employed
- Supervisory responsibility, including number typically supervised directly and through subordinates
- Relationship between work and highest degree, including reasons for employment outside the highest degree field
- Typical work activities (in 14 categories), including primary and secondary work activities
- Licensing and certification if required, recommended, or held
- U.S. government support for research, including supporting agencies or departments
- Second job, including occupation, salary, and relationship between work and highest degree field

For the unemployed and those not in the labor force:

- Reasons for not working during the reference week
- When last worked
- Job last worked

### Other Work-Related Information

- Membership in professional societies and associations, including meeting attendance
- Participation in work-related training activities, including types of training and reasons for participation

### Education

- First bachelor’s and two most recent degrees — level, degree field (major and minor), when awarded
- Earlier education — date awarded high school diploma; associate degree(s)
- Continuing education — post-degree college courses, reasons and field of study; employer financing

## Other Information

Family-related:

- Marital status
- Spouse's employment status; if working full/part-time, technical expertise required on job
- Children living at home (and ages)
- Parents' educational attainment

Demographics:

- Citizenship status (by type)
- Age
- Race/ethnicity
- Sex
- Disability
- Country of birth

## Special modules

1993: Labor force status in 1988

- Type of employer and job
- If different from current job, reasons for changing employer or job
- Years of professional work experience

1995 (SDR only): Post-doctoral experience

- Whether ever held a post-doctoral position
- Number of post-docs held over career
- Type of employer, including types of benefits offered
- Whether current job was a post-doctoral position

1995 (NSCG and SDR only): Patent and publication activity

- Number of articles or other publications authored by respondent
- Number of patent applications, patents awarded and commercializations attributed to respondent

1997: Alternative or temporary work experience

- Whether relationship to employer was alternative or temporary (consulting, contracting, etc.)
- Reasons for such work arrangements
- Whether benefits were provided, and if so, types of benefits

## HOW CAN AN ANALYST USE SESTAT?

SESTAT is useful not only for describing the characteristics of the U.S. scientists and engineers, but also for answering questions about this group. For example, it can be used to describe the educational background of engineers or social scientists. Analysts can use the system

to show what occupations people with physics degrees hold. It can also be used to answer questions such as how age or gender are related to salary distributions for scientists or engineers working in different fields or with different degrees. SESTAT is intended to support this type of research, for both policy analysis and general information.

The Division of Science Resources Studies has produced a variety of analyses that show the capability of SESTAT for analyzing characteristics and issues about the U.S. population of scientists and engineers, including the following examples:

- Two Issue Briefs, *"Degrees and Occupations in Engineering: How Much Do They Diverge?"* (NSF 99-318), and *"How Much Does the U.S. Rely on Immigrant Engineers?"* (NSF 99-327), and the topical report, *The Education and Employment of Engineering Graduates* (forthcoming), examine individuals within a certain disciplinary field or occupational category.
- The Issue Brief, *"What Follows the Postdoctorate Experience? Employment Patterns of 1993 Postdocs in 1995"* (NSF 99-307) uses data from two rounds of one of the SESTAT surveys, the SDR, to examine individuals with doctorates earned at U.S. institutions.
- SDR data was also used to explore unemployment in the Issue Brief, *"Is The Gender Gap In Unemployment Disappearing?"* (NSF 97-323).
- Salary differences among recent bachelor's and master's degree earners from the NSRCG were the subject of the Data Brief, *"Recent Engineering Graduates Out-Earn Their Science Counterparts"* (NSF 96-327).
- SESTAT has also been used to examine individuals with certain labor market characteristics. For example, recent bachelor's and master's degrees earners working in small businesses are the subject of the topical report, *Will Small Business Become the Nation's Leading Employer of Graduates with Bachelor's Degrees in Science and Engineering?* (NSF 99-322).
- The SESTAT population has been used to highlight aspects of the workforce of scientists and engineers. All individuals employed as scientists and engineers were discussed in the Data Brief,

*“Employment of Scientists and Engineers Reaches 3.2 Million in 1995”* (NSF 98-325).

The concepts used to develop the SESTAT integrated database are broad and general in order to provide the capability for flexible and customized analyses. As such, SESTAT users may tailor their inquiries and research to address a variety of issues about the population of scientists and engineers.

The SESTAT system is available to the public on the World Wide Web at <http://sestat.nsf.gov>. This site contains public use versions of the SESTAT integrated databases for 1993 and 1995, as well as a public use version of the 1993 NSCG database. The 1997 integrated database will be available in Spring 1999. The National Science Foundation ensures the confidentiality of all of the data collected and available in the SESTAT system.